УДК 512.532.2

И. А. Гольдберг

ПРОБЛЕМА КОНЕЧНОГО БАЗИСА ДЛЯ МОНОИДОВ ИНЪЕКТИВНЫХ НАПРАВЛЕННЫХ ПРЕОБРАЗОВАНИЙ*

Введение

Пусть Σ – счетный алфавит, через Σ^+ и Σ^* мы обозначим соответственно свободную полугруппу и свободный моноид над Σ. *Тождеством* над алфавитом Σ называется пара слов $u, v \in \Sigma^+$, которая обозначается посредством формального равенства $u = v$. Моноид M *удовлетворяет тождеству* $u = v$, если для любого гомоморфизма $\varphi\colon \Sigma^+ \to M$ выполняется $u\varphi = v\varphi$. Если I – некоторое множество тождеств, то говорят, что тождество $u = v$ *следует* из I, если любой моноид M, удовлетворяющий всем тождествам из I, удовлетворяет также тождеству $u = v$. Моноид M называется *конечно базируемым*, если все тождества этого моноида следуют из некоторого конечного набора его тождеств (*базиса тождеств* моноида M); в противном случае M называется *бесконечно базируемым*. Пусть $\mathcal{M}$ – некоторый класс конечных моноидов. *Проблема конечного базиса для класса* $\mathcal{M}$ состоит в том, чтобы определить, какие моноиды из $\mathcal{M}$ являются конечно базируемыми и какие – бесконечно базируемыми.

Решение проблемы конечного базиса известно для класса всех конечных групп. А именно, в работе [5] доказано, что любая конечная группа является конечно базируемой. Для конечных моноидов это условие может не выполняться. Первый пример такого рода был приведен в работе [6]. Им стал шестиэлементный моноид Брандта B_2^1, который может быть представлен в виде полугруппы 2×2-матриц

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

относительно обычного матричного умножения. Естественно возникает задача классификации конечных моноидов с точки зрения конечной базируемости их тождеств. За последние 40 лет в этом направлении проделана большая

*Работа выполнена при содействии Президентской программы поддержки ведущих научных школ (грант № НШ-2227.2003.01), программы «Университеты России – фундаментальные исследования» Министерства образования Российской Федерации (проект № УР.04.01.437), Федерального агентства по образованию (гранты № 49123 и А04-2.8-928).

работа, результаты которой по состоянию на 1985 и 2000 гг. систематизированы в обзорных статьях [1, гл. II] и [7] соответственно. Однако для ряда конкретных конечных моноидов, важных для теории и приложений, все еще неизвестно, конечен ли их базис тождеств. Исследованию некоторых таких «неподдающихся» моноидов посвящена данная работа.

Пусть X_n – n-элементное линейно упорядоченное множество, которое для определенности мы будем отождествлять с отрезком натурального ряда $1, \ldots, n$. Среди всех моноидов преобразований этого множества мы выделим так называемый *базисный каркас*, состоящий из тех моноидов, преобразования которых характеризуются некоторой комбинацией следующих четырех фундаментальных свойств:

- всюду определенность;
- инъективность;
- монотонность (частичное преобразование α называется *монотонным*, если для всех i, j, принадлежащих области определения α, из $i \leqslant j$ следует, что $i.\alpha \leqslant j.\alpha$);
- направленность (частичное преобразование α называется *направленным*, если $i \leqslant i.\alpha$ для каждого i из области определения α).

Перечислим моноиды преобразований множества X_n, которые характеризуются каждым из названных свойств по отдельности:

- $\mathcal{T}_n$, моноид всех всюду определенных преобразований, *симметрический моноид*;
- $\mathcal{I}_n$, моноид всех частичных инъективных преобразований, *симметрический инверсный моноид*;
- $\mathcal{PO}_n$, моноид всех частичных монотонных преобразований;
- $\mathcal{PE}_n$, моноид всех частичных направленных преобразований.

Каждый другой моноид из базисного каркаса получается как теоретико-множественное пересечение каких-то из только что указанных моноидов. Например, моноид $\mathcal{S}_n = \mathcal{T}_n \cap \mathcal{I}_n$ состоит из всех всюду определенных инъективных преобразований и, разумеется, есть не что иное, как группа всех перестановок множества $\{1, \ldots, n\}$ (симметрическая группа).

При $n > 2$ базисный каркас состоит из 13 моноидов, изображенных на диаграмме (рис. 1). Нетривиальные моноиды, представленные на диаграмме, интенсивно исследовались с самых разных точек зрения. Много внимания уделялось, например, вопросам их характеризации в терминах образующих и определяющих соотношений, см. недавний обзор [2]. Псевдомногообразия,

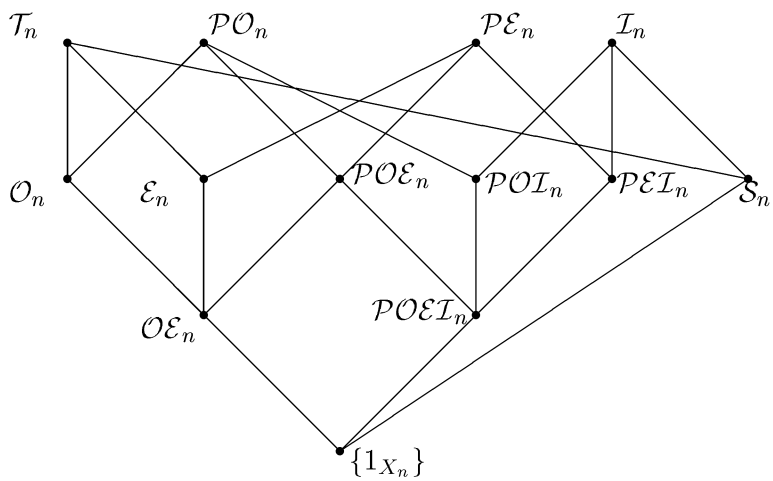

Рис. 1

порождаемые этими моноидами, оказались весьма важными с точки зрения приложений алгебры в теории формальных языков, см. обзор [4].

Однако проблема конечного базиса для ряда моноидов базисного каркаса долгое время оставалась открытой. Очевидно, что моноид $\mathcal{S}_n$ (являющийся группой) и $\{1_{X_n}\}$ конечно базируемы. С другой стороны, несложно показать, что моноиды $\mathcal{T}_n$, $\mathcal{PO}_n$, $\mathcal{I}_n$, $\mathcal{O}_n$ и $\mathcal{POI}_n$ бесконечно базируемы при $n \geqslant 3$. Проблема конечного базиса для $\mathcal{OE}_n$ была решена в [8], где было показано, что этот моноид бесконечно базируем тогда и только тогда, когда $n \geqslant 5$. Бесконечная базируемость моноидов $\mathcal{E}_{n+1}$, $\mathcal{PE}_n$ и $\mathcal{POE}_n$ при $n \geqslant 4$ была доказана в [3]. Вопрос о конечной базируемости оставшихся моноидов $\mathcal{PEI}_n$ и $\mathcal{POEI}_n$ был в явном виде сформулирован в [7].

Основным результатом данной работы является следующее утверждение.

Теорема 1. *Для любого натурального числа $n \geqslant 4$ моноиды $\mathcal{PEI}_n$ и $\mathcal{POEI}_n$ не имеют конечного базиса тождеств.*

Отметим, что ситуация, когда в последовательности моноидов преобразований конечного множества (заиндексированной в соответствии с размером этого множества) все моноиды являются бесконечно базируемыми за исключением, быть может, нескольких в начале последовательности, является весьма общей. Исключительно интересной представляется проблема, состоящая в том, чтобы отыскать причину, по которой «достаточно большие» моноиды преобразований оказываются бесконечно базируемыми. Однако имеющиеся

на сегодняшний день методы доказательства бесконечной базируемости не дают подходов к решению столь общей задачи.

Проблема конечного базиса для моноидов $\mathcal{PEI}_2$, $\mathcal{PEI}_3$, $\mathcal{POEI}_2$ и $\mathcal{POEI}_3$ остается открытой.

Работа состоит из двух параграфов. В первом параграфе приводится описание тождеств рассматриваемых моноидов преобразований. Второй параграф посвящен доказательству теоремы 1.

1. Описание тождеств моноидов $\mathcal{PEI}_n$ и $\mathcal{POEI}_n$

Введем необходимые определения. Пусть v, w – слова над алфавитом Σ. Будем говорить, что слово v является *разбросанным подсловом* слова w, если $v \equiv a_1 \cdots a_m$, где $a_1, \ldots, a_m \in \Sigma$, и можно подобрать такие слова $w_0, w_1, \ldots, w_{m-1}, w_m \in \Sigma^*$, что

$$w \equiv w_0 a_1 w_1 \cdots w_{m-1} a_m w_m; \tag{1}$$

другими словами это означает, что буквы слова v входят в w в качестве подпоследовательности. Очевидно, что если v является разбросанным подсловом слова w, то w может иметь несколько представлений в виде (1). Представление (1), для которого справедливы условия $a_i \notin c(w_{i-1})$, где $i = 1, \ldots, m$, будем называть *первым вхождением слова v в слово w*, а факторы $w_0, w_1, \ldots, w_m$, участвующие в этом представлении, – *факторами первых вхождений v в w*. Если для слова w существует всего одно представление вида (1), то слово v назовем *уникально разбросанным подсловом слова w*. Следующее свойство уникально разбросанных подслов будет полезно нам в дальнейших рассуждениях.

Предложение 1. *Пусть v – разбросанное подслово слова w и представление (1) является первым вхождением слова v в w. Для того чтобы слово v являлось уникально разбросанным подсловом слова w, необходимо и достаточно, чтобы выполнялись условия $a_i \notin c(w_i)$ для всех $i = 1, \ldots, m$.*

Доказательство. Необходимость данного условия практически очевидна. Действительно, пусть найдется такое значение i_0, что $a_{i_0} \in c(w_i)$. Следовательно, слово w_{i_0} можно представить в виде $w_{i_0} = u_{i_0} a_{i_0} u'_{i_0}$ для некоторых слов $u_{i_0}, u'_{i_0} \in \Sigma^*$. Обозначим через w'_{i_0-1} слово $w_{i_0-1} a_{i_0} u_{i_0}$. Тогда слово w может быть представлено в виде

$$w \equiv w_0 a_1 \cdots a_{i_0-1} w'_{i_0-1} a_{i_0} u'_{i_0} a_{i_0+1} \cdots a_m w_m,$$

что противоречит тому, что v является уникально разбросанным подсловом слова w.

Предположим теперь, что для всех значений $i = 1, \ldots, m$ выполняются условия $a_i \notin c(w_i)$. Покажем, что v является уникально разбросанным подсловом w. Предположим противное, т. е. пусть существует еще одно представление слова w в виде

$$w \equiv w'_0 a_1 w'_1 \cdots w'_{m-1} a_m w'_m. \tag{2}$$

Обозначим через k_i и k'_i номера позиций букв a_i в представлениях (1) и (2) соответственно. Поскольку (1) является первым вхождением слова v в w, то, очевидно, выполняются неравенства $k_i \leqslant k'_i$ для всех $i = 1, \ldots, m$. Следуя нашему предположению, найдется такое число i_0, что $k_{i_0} < k'_{i_0}$. Поскольку буква a_{i_0} не содержится в слове w_{i_0}, то $k_{i_0+1} \leqslant k'_{i_0}$, откуда $k_{i_0+1} < k'_{i_0+1}$. Индуктивно продолжая эти рассуждения, мы получим, что $k_m < k'_m$, следовательно, $a_m \in c(w_m)$, что противоречит условию предложения.

Обозначим через C_k набор всех таких тождеств $w = w'$, что выполняются следующие три условия:

1) $c(w) = c(w')$;
2) для любого $m \leqslant k$ множества уникально разбросанных подслов длины m слов w и w' должны совпадать;
3) если для уникально разбросанного подслова $v = a_1 \cdots a_m$ длины $m \leqslant k$ имеют место представления

$$w \equiv w_0 a_1 w_1 \cdots w_{m-1} a_m w_m$$

и

$$w \equiv w'_0 a_1 w'_1 \cdots w'_{m-1} a_m w'_m,$$

то для всех $i = 0, \ldots, m$ должны совпадать алфавиты факторов первых вхождений: $c(w_i) = c(w'_i)$.

Основным результатом данного параграфа является следующее утверждение.

Предложение 2. *Для любого натурального числа n множество C_n совпадает с множеством всех тождеств, выполненных в моноидах $\mathcal{PEI}_{n+1}$ и $\mathcal{POEI}_{n+1}$.*

Доказательство. Пусть M – любой из моноидов $\mathcal{PEI}_{n+1}$ и $\mathcal{POEI}_{n+1}$. Мы покажем, что тождество $w = w'$ выполняется в M тогда и только тогда, когда это тождество содержится в множестве C_n.

Пусть тождество $w = w'$ выполнено в M. Проверим, что алфавиты слов w и w' совпадают. Предположим противное, пусть $x \in c(w)$ и $x \notin c(w')$. Обозначим через 0_{n+1} такое преобразование, что $i.0_{n+1}$ не определено для любого $i = 1, \ldots, n+1$. Рассмотрим следующий гомоморфизм $\psi : \Sigma^* \to M$:

$$a\psi = \begin{cases} \varepsilon, & \text{если } a \neq x, \\ 0_{n+1}, & \text{если } a = x. \end{cases}$$

Тогда очевидно, что $w\psi = \varepsilon$, а $w'\psi = 0_{n+1}$, что противоречит выполнимости тождества $w = w'$.

Пусть $v \equiv a_1 a_2 \cdots a_m$, где $m \leqslant n$, – уникально разбросанное подслово слова w. Рассмотрим первое вхождение (1) слова v в w и определим гомоморфизм $\varphi : \Sigma^* \to M$ по следующему правилу: пусть $a \in \Sigma$, положим

$$i.(a\varphi) = \begin{cases} i, & \text{если } a \in c(w_{i-1}), \\ i+1, & \text{если } a = a_i, \\ \text{не определено}, & \text{в противном случае,} \end{cases} \qquad \text{для } i = 1, \ldots, m;$$

$$(m+1).(a\varphi) = \begin{cases} m+1, & \text{если } a \in c(w_m), \\ \text{не определено}, & \text{если } a \notin c(w_m); \end{cases}$$

и $j.(a\varphi) = j$ для $j = m+2, \ldots, n+1$. Легко понять, что определенное таким образом преобразование является направленным и монотонным. Покажем, что данное преобразование является еще и инъективным. Очевидно, что достаточно проверить, что для любого $i = 1, \ldots, n$ справедливо $i.(a\varphi) \neq (i+1).(a\varphi)$. Предположим противное: пусть число i_0 таково, что $i_0.(a\varphi) = (i_0+1).(a\varphi)$. Тогда, пользуясь определением преобразования $a\varphi$, мы можем заключить, что $i_0.(a\varphi) = (i_0+1).(a\varphi) = i_0 + 1$. Из этого следует, что $a = a_{i_0}$ и $a \in c(w_{i_0})$, что противоречит тому, что слово v является уникально разбросанным подсловом слова w. Следовательно, преобразование $a\varphi$ является инъективным и поэтому содержится в моноиде M.

По построению $1.(w\varphi) = m + 1$, следовательно, $1.(w'\varphi) = m + 1$. Легко проверить, что данное условие может быть выполнено только в том случае, если v является разбросанным подсловом w' и для первого вхождения v в w' справедливы включения $c(w'_i) \subseteq c(w_i)$, где $i = 0, 1, \ldots, m$. Из этих условий следует и тот факт, что слово v является уникально разбросанным подсловом слова w'. Обратные включения получаются из симметричных рассуждений.

Теперь предположим, что тождество $w = w'$ содержится в множестве C_n. Покажем, что для любого гомоморфизма $\varphi : \Sigma^* \to M$ и любого натурального числа $k_0 = 1, \ldots, n+1$ выполняется $k_0.(w\varphi) = k_0.(w'\varphi)$. Построим разбросан-

ное подслово $v \equiv a_1 \cdots a_m$ слова w, удовлетворяющее следующим свойствам:

$$w \equiv w_0 a_1 w_0', \quad \text{причем} \quad k.(w_0\varphi) = k,\ k.((w_0 a_1)\varphi) = k_1 \neq k;$$
$$w_0' \equiv w_1 a_2 w_1', \quad \text{причем} \quad k_1.(w_1\varphi) = k_1,\ k_1.((w_1 a_2)\varphi) = k_2 \neq k_1;$$
$$\cdots$$
$$w_{m-2}' \equiv w_{m-1} a_m w_m, \quad \text{причем} \quad k_{m-1}.(w_{m-1}\varphi) = k_{m-1};$$
$$k_{m-1}.((w_{m-1} a_m)\varphi) = k_m \neq k_{m-1};$$
$$k.(w\varphi) = k_m.$$

Очевидно, что длина построенного слова v не превосходит n, а представление

$$w \equiv w_0 a_1 w_1 a_2 \cdots w_{m-1} a_m w_m$$

есть первое вхождение слова v в w. Покажем, что v является уникально разбросанным подсловом слова w. Действительно, если предположить, что для некоторого $i = 1, \ldots, m$ справедливо условие $a_i \in c(w_i)$, то по построению w_i имеем $k_i.(w_i\varphi) = k_i$. Поскольку $a_i \in c(w_i)$, то и $k_i.(a_i\varphi) = k_i$. Но в соответствии с выбором a_i выполняется условие $k_{i-1}.(a_i\varphi) = k_i$, что противоречит инъективности $a_i\varphi$. Следовательно, слово v является уникально разбросанным подсловом слова w.

Тогда слово v является также уникально разбросанным подсловом слова w'. При этом первое вхождение слова v в w'

$$w' \equiv w_0' a_1 w_1' a_2 \cdots w_{m-1}' a_m w_m'$$

удовлетворяет условиям $c(w_0) = c(w_0')$, $c(w_1) = c(w_1')$,…,$c(w_m) = c(w_m')$. Очевидно, что

$$k_0.(w_0\varphi) = k_0.(w_0'\varphi) = k_0, \quad k_0.((w_0 a_1 w_1)\varphi) = k_0.((w_0' a_1 w_1')\varphi) = k_1, \ \ldots$$
$$\ldots\ , k_0.((w_0 a_1 w_2 \cdots w_{m-1} a_m w_m)\varphi) = k_0.((w_0' a_1 w_1' \cdots w_{m-1}' a_m w_m')\varphi) = k_m.$$

Все необходимые случаи рассмотрены, предложение доказано.

2. Доказательство теоремы 1

Данная теорема немедленно следует из предложения 2 и следующего результата.

Предложение 3. *Для любого натурального числа $n \geqslant 1$ система тождеств C_{n+2} бесконечно базируема.*

Доказательство. Для доказательства бесконечной базируемости системы тождеств C_{n+2} мы для любого достаточно большого числа m построим тождество от m переменных, содержащееся в C_{n+2}, которое не может быть выведено из тождеств этой системы от меньшего числа переменных.

При доказательстве данного предложения мы будем пользоваться теоремой Биркгофа о полноте, дающей синтаксическое описание понятия выводимости из некоторой системы тождеств. Напомним формулировку этой теоремы.

Теорема Биркгофа 1. *Нетривиальное полугрупповое тождество $w = w'$ следует из системы тождеств I тогда и только тогда, когда найдутся слова $v_0, v_1, \ldots, v_k \in \Sigma^+$ такие, что $w \equiv v_0$, $w' \equiv v_k$ и для любого $i = 1, \ldots, k$ существуют $s_i, t_i \in \Sigma^*$ и гомоморфизм $\zeta_i : \Sigma^+ \to \Sigma^+$, что $v_{i-1} \equiv s_i(u_i\zeta_i)t_i$, $v_i \equiv s_i(u'_i\zeta_i)t_i$ и тождество $u_i = u'_i$ принадлежит системе I.*

Пусть $x_1, \ldots, x_m$ – попарно различные буквы алфавита Σ. Через $\overrightarrow{z}$ и $\overleftarrow{z}$ мы будем обозначать следующие слова:

$$\overrightarrow{z} \equiv x_1 \cdots x_m, \quad \overleftarrow{z} \equiv x_m \cdots x_1.$$

Напомним, что слово $w \in \Sigma^*$ называется *изотермом относительно системы тождеств I*, если I не содержит нетривиальных тождеств, одной из частей которых является слово w. Прежде чем приступать к рассмотрению случаев, мы докажем следующий вспомогательный результат, который будет нам полезен в дальнейшем.

Лемма 1. *Пусть w, w' – слова над Σ, $\varphi : \Sigma^* \to \Sigma^*$ – гомоморфизм. Если $c(w) \supseteq c(w')$, $w\varphi \equiv \overrightarrow{z}$ и $w'\varphi \equiv \overleftarrow{z}$, то слово w' содержит как минимум m различных букв.*

Доказательство. Поскольку слово $w'\varphi$ есть произведение различных букв, то все буквы слова w' также различны. Кроме того, слова $w\varphi$ и $w'\varphi$ не имеют общих подслов длины 2. Следовательно, для любой буквы $y \in c(w')$ длина слова $y\varphi$ не превосходит 1. Поэтому слово w' содержит как минимум m различных букв.

Для любого $m \geqslant 1$ рассмотрим тождество

$$\overrightarrow{z}(\overleftarrow{z})^n x_m^2 = \overrightarrow{z}(\overleftarrow{z})^n x_m^3, \tag{3}$$

которое, очевидно, содержит m переменных. Мы покажем, что данное тождество содержится в системе C_{n+2} и не может быть выведено из тождеств этой системы от меньшего числа переменных.

Лемма 2. *При любых m тождество* (3) *содержится в системе* C_{n+2}.

Доказательство. Легко заметить, что алфавиты слов в правой и левой частях тождества (3) совпадают. Проверим, что эти слова содержат одинаковые множества уникально разбросанных подслов. Очевидно, что любое уникально разбросанное подслово v слова $\overrightarrow{z}(\overleftarrow{z})^n x_m^2$ такое, что длина слова v не превосходит $n+2$, не может заканчиваться буквой x_m. Следовательно, слово v является разбросанным подсловом слова $\overrightarrow{z}(\overleftarrow{z})^n$, а значит и слова $\overrightarrow{z}(\overleftarrow{z})^n x_m^3$. Из того что v является уникально разбросанным подсловом слова $\overrightarrow{z}(\overleftarrow{z})^n x_m^2$, немедленно следует выполнение этого условия и для слова $\overrightarrow{z}(\overleftarrow{z})^n x_m^2$. Аналогично проверяется и то, что любое уникально разбросанное подслово слова $\overrightarrow{z}(\overleftarrow{z})^n x_m^3$ является уникально разбросанным подсловом слова $\overrightarrow{z}(\overleftarrow{z})^n x_m^2$. При этом легко понять, что алфавиты факторов первых вхождений для общих уникально разбросанных подслов будут совпадать.

Лемма 3. *Пусть $w \in \Sigma^+$ – собственное подслово слова $\overrightarrow{z}(\overleftarrow{z})^n x_m^2$. Тогда w является изотермом относительно системы* C_{n+2}.

Доказательство. Очевидно, что достаточно проверить, что изотермами являются слова $w \equiv x_2 \cdots x_m(\overleftarrow{z})^n x_m^2$ и $w' \equiv \overrightarrow{z}(\overleftarrow{z})^n x_m$. Пусть тождество $w = v$ содержится в C_{n+2}. Слово $x_2 x_3^{n+1}$ является уникально разбросанным подсловом слова w. Следовательно, по определению системы C_{n+2}, это слово является уникально разбросанным подсловом слова v, а из равенства алфавитов факторов первых вхождений следует, что $v \equiv x_2 x_3 v_1$ для некоторого $v_1 \in \Sigma^*$. Рассматривая слова $x_3 x_4^{n+1}, \ldots, x_{m-2} x_{m-1}^{n+1}$, аналогичным образом доказывается, что $v \equiv x_2 x_3 \cdots x_{m-1} v_2$. Слово x_{m-1}^{n+1} является уникально разбросанным подсловом слова w, поэтому $v \equiv x_2 x_3 \cdots x_{m-1} x_m v_3$. Проведя аналогичное рассуждение для слова $x_m^2 x_{m-1}^n$, мы получим, что $v \equiv x_2 x_3 \cdots x_m x_m x_{m-1} v_4$. Далее, рассмотрев уникально разбросанное подслово $x_{m-1}^2 x_{m-2}^n$, легко понять, что $v \equiv x_2 x_3 \cdots x_m x_m x_{m-1} x_{m-2} v_5$. Продолжая этот процесс для слов вида $x_p^{k+1} x_{p-1}^{n-k+1}$ для $k = 1, \ldots, n$ и $p = m, \ldots, 2$, принимая во внимание тот факт, что длина этих слов равна $n+2$, получаем, что $v \equiv x_2 x_3 \cdots x_m(\overleftarrow{z})^n v_6$. И наконец, рассмотрев слово $x_1^n x_m^2$, получаем, что $v \equiv x_2 x_3 \cdots x_m(\overleftarrow{z})^n x_m^2 \equiv w$.

Доказательство того, что слово $\overrightarrow{z}(\overleftarrow{z})^n x_m$ является изотермом, проводится аналогично.

Лемма 4. *Пусть нетривиальное тождество $w = \overrightarrow{z}(\overleftarrow{z})^n x_m^2$ содержится в C_{n+2}. Тогда слово w имеет вид*

$$w \equiv \overrightarrow{z}(\overleftarrow{z})^n x_m^k$$

для некоторого $k > 2$.

Доказательство. Проведя рассуждения, аналогичные доказательству леммы 3, можно получить, что слово w имеет вид $w \equiv \overrightarrow{z}(\overleftarrow{z})^n v$ для некоторого слова $v \in \Sigma^+$. Предположим, что $c(v) \neq \{x_m\}$. Тогда найдется такая буква $x_j \neq x_m$, что $x_j \in c(v)$. Следовательно, буква x_j встречается в слове w более $n+1$ раза, что противоречит тому, что x_j^{n+1} является общим уникально разбросанным подсловом слов w и $\overrightarrow{z}(\overleftarrow{z})^n x_m^2$. Таким образом, $c(v) = \{x_m\}$. Поскольку слово $\overrightarrow{z}(\overleftarrow{z})^n x_m$ является изотермом относительно C_{n+2}, то слово w может быть представлено в виде $w \equiv \overrightarrow{z}(\overleftarrow{z})^n x_m^k$ для некоторого $k > 2$.

Нам необходимо проверить, что тождество $\overrightarrow{z}(\overleftarrow{z})^n x_m^2 = \overrightarrow{z}(\overleftarrow{z})^n x_m^3$ не может быть выведено из тождеств моноидов $\mathcal{PEI}_{n+3}$ от меньшего числа переменных. Для этого рассмотрим произвольный вывод данного тождества, т. е. последовательность

$$\overrightarrow{z}(\overleftarrow{z})^n x_m^2 \equiv v_0, v_1, \ldots, v_r \overrightarrow{z}(\overleftarrow{z})^n \equiv \overrightarrow{z}(\overleftarrow{z})^n x_m^3,$$

где для любого i найдутся слова $s_i, t_i \in \Sigma^*$, гомоморфизм $\zeta_i : \Sigma^+ \to \Sigma^+$ и нетривиальное тождество $u_i = u_i'$, содержащееся в C_{n+2}, что

$$v_{i-1} \equiv s_i(u_i\zeta_i)t_i, \quad v_i \equiv s_i(u_i'\zeta_i)t_i.$$

Из леммы 3 следует, что $s_1 \equiv t_1 \equiv \varepsilon$, т. е. $\overrightarrow{z}(\overleftarrow{z})^n x_m \equiv \zeta_1 u_1$, а по лемме 4 слово $u_1'\zeta_1$ имеет вид $u_1'\zeta_1 \equiv \overrightarrow{z}(\overleftarrow{z})^n x_m^k$ для некоторого $k > 2$.

Поскольку слова u_1 и u_1' не равны и $u_1\zeta_1$ является подсловом $u_1'\zeta_1$, то возможны только следующие два варианта: либо u_1 и u_1' различаются хотя бы в одной позиции, либо u_1 является подсловом слова u_1'. Предположим, что имеет место первый вариант, т. е.

$$u_1 = s_0 a s, \quad u_1' = s_0 b s', \tag{4}$$

где $p \in \Sigma^*$, $a, b \in \Sigma$, $a \neq b$. При этом можно считать, что $a\zeta_1 \neq b\zeta_1$.

Пусть, как и выше, преобразование τ переставляет буквы a и b: $a\tau = b$, $b\tau = a$. Справедлива следующая лемма.

Лемма 5. *Пусть для некоторого $\ell \geqslant 0$ имеет место представление*

$$u_1 \equiv s_0 a s_1 b s_2 \cdots a\tau^\ell s_{\ell+1}, \quad u_1' \equiv s_0 b s_1' a s_2' \cdots b\tau^\ell s_{\ell+1}',$$

где $s_0, s_1, \ldots, s_{\ell+1}, s_1', \ldots s_{\ell+1}' \in \Sigma^*$, $a, b \notin c(s_i)$, $a, b \notin c(s_i')$ для всех $0 < i \leqslant \ell$. Тогда найдутся такие слова $\overline{s}_0, \ldots, \overline{s}_{\ell+1}, \overline{s}_{\ell+2}, \overline{s'}_1, \ldots, \overline{s'}_{\ell+1}, \overline{s'}_{\ell+2}$, что

$$u_1 \equiv \overline{s}_0 a \overline{s}_1 b \overline{s}_2 \cdots a\tau^{\ell+1}\overline{s}_{\ell+2}, \quad u_1' \equiv \overline{s}_0 b \overline{s'}_1 a \overline{s'}_2 \cdots b\tau^{\ell+1}\overline{s'}_{\ell+2},$$

где $a, b \notin c(\overline{s}_i)$, $a, b \notin c(\overline{s'}_i)$ для всех $0 < i \leqslant \ell + 1$.

Доказательство. Обозначим через p и q количество вхождений букв a и b соответственно в слово u_1. Поскольку тождество $u_1 = u_1'$ содержится в C_{n+2}, то $c(u_1) = c(u_1')$. В частности, буква b содержится в слове u_1. Нетрудно понять, что слова $a\zeta_1$ и $b\zeta_1$ начинаются с одной и той же буквы, следовательно сумма $p + q$ не может быть больше $n + 3$. Это, в свою очередь, означает, что $p \leqslant n + 2$ и $q \leqslant n + 2$. Следовательно, слова a^p и b^q являются уникально разбросанными подсловами слова u_1, а значит и слова u_1'. Поэтому количество вхождений букв a и b в слово u_1' также равно p и q соответственно.

Из этого следует, что для любого натурального числа $r = 1, \ldots, p-1$ алфавиты факторов слов u_1 и u_1', «зажатых» между r-м и $(r+1)$-м по порядку вхождениями буквы a, совпадают. Аналогичное условие можно сформулировать и для факторов слов u_1 и u_1', «зажатых» между последовательными вхождениями буквы b.

Кроме того, нетрудно понять, что алфавиты слов $s_\ell a\tau^\ell s_{\ell+1}$ и $s'_{\ell+1}$ совпадают. Это означает, что буква $a\tau^\ell = b\tau^{\ell+1}$ содержится в слове $s'_{\ell+1}$. Следовательно, $s'_{\ell+1} \equiv \overline{s'}_{\ell+1} b\tau^{\ell+1}\overline{s'}_{\ell+2}$, где $b\tau^{\ell+1} \notin c(\overline{s'}_{\ell+1})$. Аналогичным образом можно получить следующее представление: $s_{\ell+1} \equiv \overline{s}_{\ell+1} a\tau^{\ell+1}\overline{s}_{\ell+2}$, где $a\tau^{\ell+1} \notin c(\overline{s}_{\ell+1})$. Осталось проверить, что $a\tau^{\ell+1} \notin c(\overline{s'}_{\ell+1})$ и $b\tau^{\ell+1} \notin c(\overline{s}_{\ell+1})$.

Предположим противное, пусть $a\tau^{\ell+1} \in c(\overline{s'}_{\ell+1})$. Обозначим через r' порядковый номер вхождения буквы $b\tau^\ell$, встречающейся в слове u_1' между факторами s'_ℓ и $s'_{\ell+1}$. Очевидно, что r'-е вхождение этой буквы в слово u_1 находится перед фактором s_ℓ. Следуя нашему предположению, между r'-м и $(r'+1)$-м вхождением буквы $b\tau^\ell$ в слово u_1' присутствует буква $a\tau^{\ell+1}$. Однако это не верно для слова u_1: фактор, «зажатый» между r'-м и $(r'+1)$-м вхождением буквы $b\tau^\ell$, равен $s_\ell a\tau^\ell \overline{s}_{\ell+1}$. По условию леммы $a\tau^{\ell+1} \notin c(s_\ell)$; выше мы показали, что $a\tau^{\ell+1} \notin c(\overline{s}_{\ell+1})$. Таким образом, наше предположение неверно и $a\tau^{\ell+1} \notin c(\overline{s'}_{\ell+1})$. Условие $b\tau^{\ell+1} \notin c(\overline{s}_{\ell+1})$ проверяется аналогично.

В дальнейшем нам будет полезно следующее практически очевидное наблюдение.

Лемма 6. *Буквы a и b встречаются в слове s_0.*

Доказательство. Обозначим через t и t' количество вхождений букв a и b соответственно в слово s_0. Тогда между t-м и $(t+1)$-м вхождением буквы a в слово u_1' встречается буква b. Следовательно, b встречается и между соответствующими вхождениями буквы a в u_1, и поэтому $b \in c(s_0)$. Аналогично доказывается, что $a \in c(s_0)$.

Обозначим через x_j первую букву слов $a\zeta_1$ и $b\zeta_1$. Мы предполагали выше, что слова u_1 и u_1' могут быть представлены в виде (4). Поэтому мы можем

применить лемму 5 для $\ell = 0$ и далее по индукции применять ее до $\ell = n+1$. Тогда мы получаем, что буква x_j встретится в слове $(a\overline{s}_1 b\overline{s}_2 \cdots a\tau^{n+2}\overline{s}_{n+3})\zeta_1$ как минимум $n+3$ раза, следовательно, она не может встретиться в слове $s_0\zeta_1$, поэтому $a \notin c(s_0)$, что противоречит лемме 6.

Таким образом, мы получаем, что слово u_1 является подсловом слова u_1', т. е. найдутся такие $t \in \Sigma$, $\overline{u}_1 \in \Sigma^*$, что $u_1' \equiv u_1 t \overline{u}_1$. Несложно понять, что буква t должна содержаться в слове u_1, кроме того, количество ее вхождений не может быть меньше $n+3$. Поскольку единственным подсловом слова $\overrightarrow{z}(\overleftarrow{z})^n x_m^2$, встречающимся не менее $n+3$ раз, является x_m, то буква t должна встречаться в u_1 ровно $n+3$ раза и $t\zeta_1 \equiv x_m$. Тогда слова u_1 и u_1' могут быть представлены в виде

$$u_1 \equiv r_0 t t r_2 \cdots r_{n+1} t^2, \quad u_1 \equiv r_0 t t r_2 \cdots r_{n+1} t^2 t \overline{u}_1.$$

При этом $r_0\zeta_1 \equiv x_1 \cdots x_{m-1}$ и $r_2\zeta_1 \equiv x_{m-1} \cdots x_1$.

Покажем, что $c(r_0) \supseteq c(r_2)$. Предположим противное, т. е. пусть найдется некоторая буква x, такая что $x \in c(r_2)$, но $x \notin c(r_0)$. Поскольку $r_2\zeta_1$ есть произведение различных букв, то все буквы слова r_2 также различны. Следовательно, буква x встречается в слове r_2 ровно один раз. Тогда слово xt^{n+1} является уникально разбросанным подсловом слова u_1, но не является таковым для u_1', что противоречит определению системы C_{n+2}. Следовательно, $c(r_0) \supseteq c(r_2)$.

Применив лемму 1, мы получаем, что $|c(r_2)| = m-1$, поэтому тождество $u_1 = u_1'$ содержит как минимум m переменных.

Литература

1. Шеврин Л. Н., Волков М. В. Тождества полугрупп // Изв. вузов. Математика. 1985. № 11. С. 3–47.

2. Fernandes V. H. Presentations for some monoids of partial transformations on a finite chain: a survey // Semigroups, Algorithms, Automata and Languages / Eds. Gracinda M. S. Gomes, Pin J., Silva P. V. World Scientific, Singapore, 2002. P. 363–378.

3. Goldberg I. A. On the finite basis problem for the monoids of extensive transformations // Proc. Intern. Conf. on Semigroups and Languages. World Scientific, Singapore, 2006.

4. Higgins P. M. Pseudovarietes generated by transformation semigroups // Semigroups and Their Applications, Including Semigroup Rings / Eds. Kublanovsky S., Mihalev A., Ponizovkii J., Higgins P. St. Petersburg State Tech. Univ., St. Petersburg, 2006. P. 85–94.

5. Oates S., Powell M. B. Identical relations in finite groups // J. Algebra. 1964. Vol. 1, № 1. P. 11–39.

6. Perkins P. Bases for equational theories of semigroups // J. Algebra. 1969. Vol. 11, № 2. P. 298–314.

7. Volkov M. V. The finite basis problem for finite semigroups // Mathematica Japonica. 2001. Vol. 53, № 1. P. 171–199.

8. Volkov M. V. Reflexive relations, extensive transformations and piecewise testable languages of a given height // Internat. J. Algebra Comput. 2004. Vol. 14, № 5–6. P. 817–827.